The Analysis of Verbal Behavior 1990, 8, 1-11

An Instance of Spurious Equivalence Relations

Gregory Stikeleather
Palo Alto, California

AND

Murray Sidman
New England Center for Autism

Four normal children learned conditional discriminations that had upper-case or lower-case Greek letters as comparison stimuli, and dictated letter names as samples. Experimental stimuli were three pairs of letters; within each pair, an upper- and a lower-case letter were conditionally related to the same dictated sample. Four control stimuli, also upper- and lower-case letters, were each conditionally related to a different dictated sample. Conditional-discrimination tests for equivalence used the upper- and lower-case letters both as samples and comparisons. Untaught conditional relations between the upper- and lower-case members of each experimental stimulus pair were expected to emerge on the basis of their previously established relations to a common sample. The emergence of conditional relations between control stimuli, however, would have suggested an artifact. In test trials with the experimental stimuli as samples and comparisons, new conditional discriminations emerged as expected with all four children. With two of the children, however, consistent discriminations also emerged between control stimuli. Evidence suggested that uncontrolled features of the program for teaching the children the baseline conditional discriminations might have been responsible for the emergence of untaught conditional relations.

When citing data to support or challenge a descriptive or theoretical account, one can often prevent fruitless controversy by first making sure that the subjects' behavior was relevant to the phenomenon under discussion. Before attempting to use particular data to support or refute an account of equivalence relations, for example, one should be reasonably certain that equivalence, or the absence of equivalence, was really an outcome of the experimental operations. The apparent emergence of equivalence relations from conditional discriminations, and apparent failures of equivalence to emerge, can sometimes be explained by procedural rather than theoretical considerations.

For example, a subject's seeming failure to show equivalence between sample and comparison stimuli in a two-comparison conditional discrimination can arise from an experimenter's mistaken assumption about the nature of the conditional control. The usual assumption is that the subject's choice on each trial is controlled by a relation between the sample and what the experimenter has designated as the positive comparison, a relation that is often called "S-plus control" (Carter & Werner, 1978). S-plus control leads to equivalence between the sample and its positive comparison. The procedure may, however, have generated "S-minus control," a conditional relation that can bring about equivalence between the sample and the negative comparison. Because the actual controlling relation goes unrecognized, the subject is judged to have selected the "wrong" stim-

This reserach, carried out at Northeastern University, was part of a study conducted by the first author in partial fulfillment of requirements for a Master's degree in Psychology. Results were presented at the 1979 meeting of the Association of Behavior Analysis, Dearborn. We thank Leanna Standish and Charles Hamad for their comments on an earlier version of the manuscript, and to F. Garth Fletcher for equipment construction and maintenance.
Reprint requests should be sent to Murray Sidman, New England Center for Autism, 33 Turnpike Road, Southborough, Massachusetts 01772.

uli on the equivalence test; S-minus control generates a score of zero in equivalence tests (Carrigan, 1986; Johnson & Sidman, 1990). If the theory at issue required that equivalence fail to develop in that particular experiment, the investigator would mistakenly take this result as a confirmation of the theory.

Another potential pitfall in interpretation can be brought about by contextual control (Bush, Sidman, & de Rose, 1989; Sidman, 1986; Wulfert & Hayes, 1988). Any stimulus can be a member of more than one class, and features of the environment can determine which equivalence class a stimulus belongs to at the moment. An investigator who overlooks the possibility of contextual control may conclude mistakenly that equivalence relations between particular stimuli do not exist.

A subject's failure to show equivalence can also arise from unintended identity relations—physical resemblances between stimuli—that can override other equivalence relations. In an equivalence test, stimuli that look alike may be classed together in spite of the subjects having been taught conditional discriminations that were designed to place the stimuli in separate classes (Barnes, 1990).

Faulty specification of the controlling stimuli will inevitably cause negative results in subsequent tests of transfer of control to other stimuli (e.g., Constantine, 1981; Iversen, Sidman, & Carrigan, 1986). And even before one tests for equivalence, insufficiently rigorous criteria for the acquisition of the basic conditional discriminations can cause a failure to establish the necessary prerequisites—a failure that is likely to go unrecognized. In a two-comparison conditional discrimination, for example, a subject can achieve an accuracy score of 75 percent—seemingly significantly different from chance—even though learning completely different conditional discriminations than those the experimenter has designated and measured (Sidman, 1980).

Procedural artifacts can also generate "false positives," tests that seem to demonstrate equivalence even though equivalence does not exist. Perhaps even less likely to be detected are instances in which equivalence does exist, but for reasons unrelated to the experimental variables. Again, such data may be taken to support a particular descriptive or theoretical account when, in fact, the equivalence test was positive for reasons extraneous to the account.

For example, when procedures designed to establish equivalence relations have failed to do so, false positives may be generated if the tests alone can teach the subject the relations being tested for. It is now well established that learning does take place even during tests that omit reinforcement (e.g., Lazar, Davis-Lang, & Sanchez, 1984; Saunders, Wachter, & Spradlin, 1988; Sidman, Kirk, & Willson-Morris, 1985; Sidman, Willson-Morris, & Kirk, 1986; Sigurdardottir, Green, & Saunders, 1990; Spradlin, Cotter, & Baxley, 1973). Furthermore, subjects can learn new conditional discriminations in the absence of reinforcement without having previously learned related discriminations (Harrison & Green, 1990; Saunders, Saunders, Kirby, & Spradlin, 1988). (For more complete discussions of the conditions under which learning may take place during unreinforced conditional-discrimination tests, see Sidman, in press; Sidman, Kirk, & Willson-Morris, 1985.) If an equivalence test can by itself teach a subject the very conditional discriminations that are needed to demonstrate equivalence, then any use of that test to support an account of equivalence that is based on previous learning will be invalid.

Also, even when experimental procedures that are designed to generate equivalence fail to do so, tests might yield seemingly positive results simply because the tested stimuli are physically or conceptually similar. Such similarities are not always apparent to the experimenter, and not all subjects will react to them.

To determine whether evidence for equivalence is spurious, the ideal control would include all features present in the experimental condition, while excluding the independent variable and any features not present in the experimental condition. We report here an attempt to set up just such a control in the context of a procedure used by Sidman (1971) to demonstrate the emergence of equivalence relations from conditional discriminations. Although we did replicate the original demonstration, we also discovered possible sources of artifacts. As the literature on stimulus equivalence expands, the importance of such discoveries grows; making them known will help increase the likelihood that data cited in support of a theoretical position are first examined closely for their validity.

METHOD

The general plan of the experiment was the following: Children were taught conditional discriminations with dictated letter names as samples, and with five upper-case or five lower-case Greek letters as comparisons. As Figure 1 indicates, the children were to relate three of the dictated-name samples ("xi," "gamma," and "lambda") both to an upper- and a lower-case comparison letter. These letters were the experimental stimuli. It was expected that the children would, when tested, match the upper- to the lower-case form of each letter without having been directly taught to do so (e.g., Sidman, Cresson, & Willson-Morris, 1974). For example, with upper-case xi as the sample and five lower-case letters as comparisons, lower-case xi was the child's predicted choice; with lower-case xi as the sample and five upper-case letters as comparisons, upper-case xi was the predicted choice.

With the dictated samples "omega" and "phi," the children were taught to choose upper-case omega and phi, respectively, and with the samples "sigma" and "delta," to select lower-case sigma and delta, respectively. Upper-case omega and phi, and lower-case sigma and delta were the control stimuli. Since each was related to a different sample, it was expected that later testing would yield unpredictable conditional relations, if any, between these stimuli.

Having learned the baseline conditional discriminations, the children were tested for the emergence of new conditional dis-

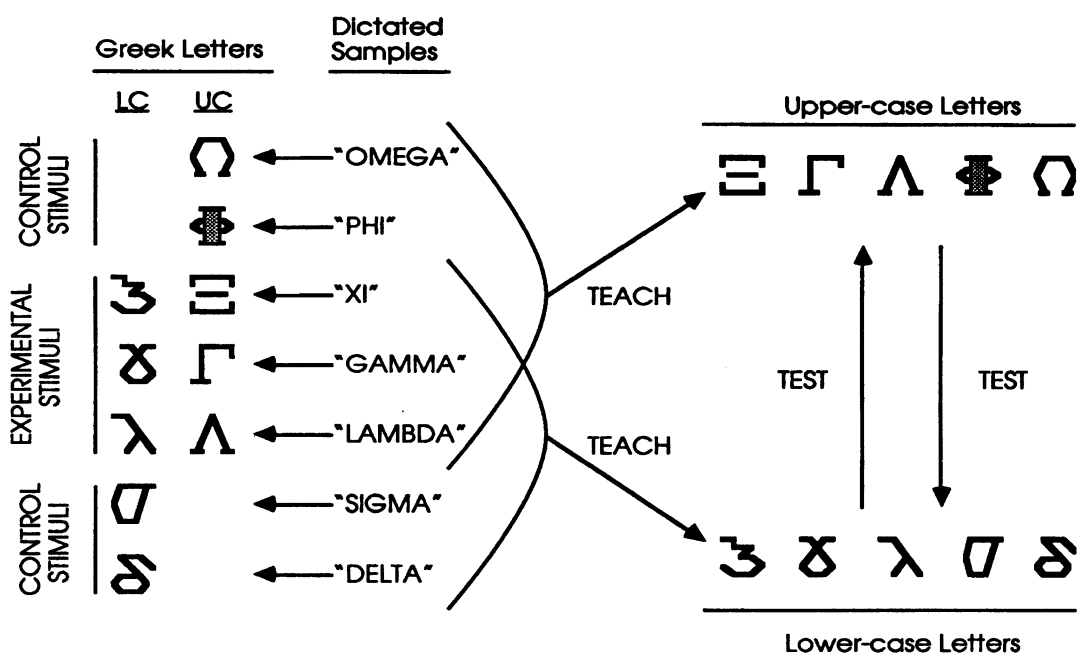

Fig. 1. Subjects were first taught baseline conditional discriminations: Each of three dictated Greek letter name samples ("xi," "gamma," and "lambda") was related to an upper-case (UC) and a lower-case (LC) comparison. Two samples ("omega" and "phi") were related only to UC, and two ("sigma" and "delta") only to LC comparisons. Subjects were then tested for the emergence of new conditional discriminations with UC samples and LC comparisons, and vice versa.

criminations in which the stimuli were upper- and lower-case letters. One set of tests used each upper-case letter as a sample, with all five lower-case letters as comparisons. The other set of tests used lower-case samples and the upper-case letters as comparisons.

The children received the same kinds of teaching experiences with all stimuli, with one exception: The two members of each pair of experimental stimuli had a history of being related to a common dictated name; the control stimuli lacked this history. It was expected, therefore, that in the tests the children would match the upper- and lower-case member of each experimental stimulus pair to each other, but would not do so with the control stimuli.

Apparatus and basic procedures have been reported elsewhere (Sidman & Tailby, 1982), and will only be summarized here. Features most relevant to the purposes of this report will be described in detail.

Subjects

Four normal children participated: two girls, Subjects S.D. and G.H., and two boys, Subjects J.P. and B.N. Their ages (years-months) at the time of their final tests were 5-5 (J.P.), 5-6 (S.D.), 6-4 (B.N.), and 6-5 (G.H.). Experimental sessions, held several days a week at the same time each day, lasted between 30 and 60 minutes.

Apparatus and Conditional-discrimination Procedures

The subject sat in front of a vertical panel that held a circle of eight translucent keys, with a ninth key in the center. Visual stimuli were projected onto the back of the keys, and auditory stimuli were dictated via a tape-and-speaker system. Pennies were dropped into a tray beside the keys, and doorbell chimes rang each time a penny was delivered. Subjects' comments were recorded on tape.

Visual sample stimuli were presented on the center key. When the sample was auditory, the center key was illuminated with white light. Comparison stimuli, always visual, were presented on the outer keys, but only the lower five of those keys were used. The maximum number of comparison stimuli presented was five, the minimum, two (see below). Comparison keys not illuminated with visual stimuli remained dark.

A trial always began with the presentation of a sample stimulus. When the subject pressed the sample key, comparisons appeared on the outer keys while the sample remained present (auditory samples continued to be repeated). The trial ended after the subject pressed a comparison key. Except when the probability of reinforcement was less than 1.00 (see below), chimes and a penny followed if the subject pressed the comparison that was programmed as correct for that sample. If the subject pressed any other comparison key, an incorrect choice was recorded and no chimes or pennies were delivered. When the subject pressed a comparison key, all stimuli disappeared. The next sample was presented after a 1.5 to 2.0-s intertrial interval.

Pre-Experimental Teaching

Subjects received no oral or written instructions. To acquaint them with the apparatus and procedures, they were given four- or five-choice conditional discriminations with familiar colors and color names. After delivering two or three pennies to accomplish magazine training, the experimenter manually demonstrated the first few conditional-discrimination trials in which the sample and correct comparison were the same color. All subjects proved able to select the appropriate comparison in the presence of a color or dictated color-name sample.

To ensure that the subjects could discriminate the stimuli to be used later, they were then given five-choice conditional discriminations in which the correct comparison was an upper-case or lower-case Greek letter identical to the sample. All subjects proved able to do this, and completed the pre-experimental teaching in one to five sessions.

Teaching

With each of five dictated Greek letter names as samples, the children learned to select the corresponding upper-case letter, and with each of five letter-name samples, they learned to select the corresponding lower-case letter. Procedurally, as shown in Figure 1, three dictated names were conditionally related both to upper- and lower-case letters, two were related only to upper-case letters, and two only to lower-case letters. The ten conditional discriminations were taught in a series of steps.

First, each subject was taught a two-comparison conditional discrimination. For example, with dictated "gamma" or "xi" as the sample, and the corresponding upper-case letters as comparisons, subjects quickly learned to select upper-case gamma when "gamma" was dictated, and upper-case xi when "xi" was dictated.

Then, as the subject mastered each conditional discrimination, a new sample and comparison stimulus were added until the subject had learned all ten conditional discriminations. The comparison stimuli on any trial were either upper or lower case, never both.

The teaching sequence varied among subjects. Table 1 shows how new conditional discriminations were introduced for each subject. In sequences I and II, the subject was not given the three-comparison task (gamma, xi, lambda) until each of the three two-comparison components (gamma, xi; gamma, lambda; xi, lambda) had been learned. In sequences III and IV, lambda was added immediately after the subject had learned the first two-comparison (gamma, xi) conditional discrimination. Subjects G.H. and B.N. differed from the others in having one or both of the abbreviated teaching sequences. The learning criterion at each teaching step was no more than one error in a set of trials that included at least five and sometimes ten trials with each sample.

After a subject had learned each set of five-comparison conditional discrimina-

Table 1

Sequences used to teach subjects conditional discriminations with upper-and lower-case Greek-letter comparisons and dictated letter name samples. Upper- and lower-case Greek letters are denoted by the first letters of their English names. Experimental stimuli: G, g, X, x, L, l; Control stimuli: P, O, s, d.

Subject	Teaching Sequence
J. P.	I, II
S.D.	I, II
G.H.	I, IV
B.N.	III, IV

Teaching Sequences (Dictated Letter-Name Samples)

Sequence	I Upper-case	II Lower-case	III Upper-case	IV Lower-case
2-choice	G X	g x	G X	g x
2-choice	G L	g l	—	—
2-choice	X L	x l	—	—
3-choice	G X L	g x l	G X L	g x l
4-choice	G X L P	g x l s	G X L P	g x l s
5-choice	G X L P O	g x l s d	G X L P O	g x l s d

tions, one set with upper- and the other with lower-case comparisons, both sets were mixed within 20-trial blocks. A subject's training was complete when at least 19 trials out of a 20-trial block were correct. The total number of teaching sessions varied from 7 for Subject G.H. to 21 for Subject J.P.

Testing

The subjects' final performance, ten conditional discriminations with dictated letter names as samples and five upper- or lower-case letters as comparisons, constituted the baseline trials in the tests. On test trials, upper-case samples were presented with lower-case comparisons, and lower-case samples with upper-case comparisons. These new five-comparison test tasks, which the subjects had never seen before, were inserted as probe trials among the baseline trials.

Probe trials were never followed by chimes or pennies. To help prevent discriminated extinction on test trials, the probability of reinforcement after correct baseline trials was reduced over several sessions to 0.20 before testing began. Subject J.P. differed from the others in having his baseline reinforcement probability reduced to 0.00.

Each 75-trial test contained 50 baseline trials and 25 probes. To maintain the overall reinforcement probability at 0.20 (no reinforcement followed probe trials), the baseline probability was increased to 0.30 during tests (for all but Subject J.P.). At the end of test sessions (and the earlier sessions in which the reinforcement probability was reduced), the children were given "identity matching" trials with colors or Greek letters to make up for reinforcements they had missed.

The complete test battery contained four 75-trial tests, two with upper-case samples and lower-case comparisons, and two with lower-case samples and upper-case comparisons. All children had the test battery at least once, starting with upper-case samples and lower-case comparisons.

RESULTS

The matrices in Table 2 summarize the test results for each subject. Within each matrix, rows designate sample stimuli, and columns designate comparison stimuli. The left-hand matrices show the percentage of probe trials on which subjects selected each lower-case comparison when each upper-case letter was a sample; the right-hand matrices show selections of upper-case comparisons when the samples were lower-case letters.

Each matrix is divided into quadrants: The upper left quadrants show the percentage of trials on which subjects selected experimental comparison stimuli when the samples were also experimental stimuli. High frequencies in the three outlined cells of each upper left quadrant would indicate the emergence of conditional discriminations that were indicative of equivalence relations between the upper- and lower-case member of each experimental stimulus pair.

The lower right quadrants show the percentage of trials on which subjects selected control comparison stimuli when the samples were also control stimuli. The lower left and upper right quadrants show how likely the subjects were to select a comparison stimulus from one set (experimental or control) when the sample was a stimulus from the other set.

The two uppermost matrices summarize Subject J.P.'s performance. When an experimental stimulus, upper-case xi, gamma, or lambda, was the sample (left-hand matrix), the child most often chose lower-case xi, gamma, or lambda, respectively. And with lower-case xi, gamma, or lambda as the sample (right-hand matrix), he most often chose the corresponding upper-case letter. These emergent conditional discriminations were symmetric; within each pair of experimental stimuli it mattered little which was the sample and which the comparison.

The child also showed comparison preferences when the sample was one of the control stimuli, upper-case phi or omega (left-hand matrix), or lower-case sigma or delta (right-hand matrix), but these preferences were not as pronounced as they were when the samples were experimental stimuli. Also, emergent conditional discriminations involving control stimuli were not symmetric. Although the subject most often selected lower-case sigma when upper-case omega was the sample, his most likely choice with lower-case sigma as the sample was not upper-case omega but upper-case phi. He also selected upper-case phi frequently when lower-case delta was the sample. With upper-case phi as the sample, however, he almost never chose lower-case sigma or delta, most often selecting lower-case gamma instead.

Subject S.D.'s tests are summarized below those of Subject J.P. Again, the results with the three pairs of experimental stimuli were consistent with the formation of equivalence relations; within each of the experimental upper- and lower-case letter pairs, symmetric conditional discriminations emerged. When a control stimulus was the sample, however, the child displayed no clear-cut comparison preferences. (Instead, she exhibited a position preference, pressing the bottom comparison key on almost 75% of the probe trials that had a control stimulus as the sample.)

Even more strongly than the first two children, Subjects G.H. and B.N. showed the emergent symmetric conditional discriminations that were required to document equivalence relations between upper- and lower-case experimental stimuli. These subjects, however, also demonstrated consistent and symmetric conditional discriminations on trials that had control stimuli as samples. With control stimuli, both children showed the same conditional discrim-

Table 2

Percentage of test trials on which subjects related each sample to each comparison. Subjects are identified in the leftmost column. Greek-letter stimuli are identified by their abbreviated names.

	UPPER-CASE SAMPLES	LOWER-CASE COMPARISONS					LOWER-CASE SAMPLES	UPPER-CASE COMPARISONS				
		xi	gam	lam	sig	del		XI	GAM	LAM	PHI	OME
J. P.	XI	93		7			xi	93		7		
	GAM		80	13	7		gam		93	7		
	LAM		20	80			lam	7	7	87		
	PHI	33	60			7	sig		20	13	53	13
	OME	13		7	53	27	del			20	67	13
S. D.	XI	100					xi	80	20			
	GAM		100				gam	30	60		10	
	LAM	7	7	87			lam	10		70	10	10
	PHI	27	13	33	20	7	sig	10	20	20	30	20
	OME	13		20	33	33	del	20		30	20	30
G. H.	XI	100					xi	100				
	GAM	10	90				gam		100			
	LAM			100			lam			100		
	PHI		10		80	10	sig				100	
	OME			10	20	70	del					100
B. N.	XI	100					xi	100				
	GAM		100				gam	10	90			
	LAM			100			lam			100		
	PHI			10	90		sig		10		90	
	OME				20	80	del					100

inations: When upper-case phi was the sample, they most often selected lower-case sigma, and with lower-case sigma as the sample, they almost always chose upper-case phi. Symmetric conditional discriminations also emerged when upper-case omega or lower-case delta was the sample.

When the sample was an experimental stimulus, the children rarely chose a control comparison. When the sample was a control stimulus, however, Subjects J.P. and S.D. selected an experimental comparison about as often as they selected a control comparison. In contrast, when the sample was a control stimulus, Subjects G.H. and B.N. almost always selected a control comparison.

DISCUSSION

Subjects J.P. and S.D. showed consistent and symmetric conditional discriminations, not directly taught, when the samples in test trials were upper- or lower-case experimental stimuli, but not when the samples were control stimuli. The difference between experimental and control stimuli indicated that the positive equivalence tests were not spurious. The children had learned to relate each member of an upper- and lower-case pair of experimental stimuli to the same dictated letter name, but to relate every control stimulus to a different name.

This different treatment—some members of one set of comparison stimuli being related to a common sample, and each member of the other set being related to a different sample—was the experimental variable. If some feature of the teaching or testing situation other than the experimental variable were sufficient for the emergence of consistent relations between upper- and lower-case experimental stimuli, similar relations would have been expected to emerge between control stimuli also.

The test performances of Subjects G.H. and B.N., however, challenged these conclusions. Both children showed symmetric relations emerging not only between experimental stimuli, but between upper- and lower-case control stimuli also, thereby calling into question the nature of the relations between experimental stimuli. For these two subjects, some artifact had to be responsible for the seeming emergence of equivalence relations.

A post-experimental interview with Subject B.N. provided clues to one possible variable. The child was given ten probe trials, with the experimenter manually controlling the intertrial intervals so that questions could be asked and answered after each trial. The following are some of the more significant excerpts from the interview:

E: How did you know to pick that one?

S: I don't kn—, well, I used my brain. Everybody has one.

E: Well, what made you pick that one instead of the others?

S: I thought which one is right.

E: How did you know it was right?

S: Well, because I both learned them at the same time. I mean, like let's say it was third. Well, it had one in the middle and it was the third—let's say it was in the middle and it was the third one I learned. So I look at the other one that I learned third on the other team.

E: O.K., let's try some more....What made you pick that one?

S: Every time? Well, because it was before the last one, and the other—and the one on the other team was also before the last one.

E: What is the last one?

S: The last one, there's two last ones, there's two of it. The two last ones are omega and delta. Want to hear the two first ones?

E: Yeah.

S: Uh, two gammas, but they both look different. One looks like a Y and one looks like an upside-down L.

E: O.K., let's try some more....Now how did you know to do that one?

S: I know it, but I told you before. I—look, they both go first.

E: Go first?

S: Right. I both learned them first.

E: And what did you learn second?

S: Second I learned the two xis. And third, I learned the two lambdas. Now it's not two of the same thing anymore.

E: So what is it?

S: Well, ther— laa—. Well, I don't remember. I think it was, let me see, one, two, three, oh, so there's two more. Another one is, uh—uh—umm, sigma and phi. And the very, very, very last one is delta and omega. And that's how I know it was that one, and don't ask me it anymore, please.

E:....So how do you know which ones to put together?

S: Well, whichever ones—like if it, if it flashed the first one I learned, I mean like gamma, then I press the other kind of gamma. The same with all of them. I told you those.

E: And the same with all of them?

S: Yeah. Yeah. Xi goes to the other xi, and lambda goes to the other lambda. And all the rest goes to—the one I learned first goes to the one I learned first. The one I learned second is the one, and then the answer is the one I learned second. And it goes on and on and on and until it gets up to one, two, three, four, five times. And if you count the both teams together, it's one, two, three, four, five, six, seven, eight, nine, ten.

Subject B.N.'s remarks called attention to the teaching program. In all teaching sequences (Table 1), conditional discriminations involving auditory samples and upper-case comparisons were taught before those involving lower-case comparisons. This could have been the basis for the subject's classification of the stimuli into two "teams."

Also, omega was always introduced as the fifth upper-case letter, and delta as the fifth lower-case letter. Subject B.N. reported matching upper-case omega to lower-case delta because both stimuli had been taught last in their respective teams. Upper-case phi and lower-case sigma were the fourth letters to be introduced, and the subject gave that reason for matching them.

The teaching sequences that Subjects G.H. and B.N. underwent differed from those of the other subjects. Subject G.H. was not questioned about her performance. Perhaps, however, the direct change from a two-comparison to a three-comparison conditional discrimination in sequences III and/or IV highlighted lambda's status as the third stimulus introduced, and directed the subjects' attention to the sequence of stimulus introduction. This variable might account for the differences between the two pairs of subjects.

Subject B.N.'s comments about the experimental stimuli suggested multiple control. He reported matching the experimental stimuli not only because, within their teams, xi, gamma, and lambda occupied corresponding positions in the teaching sequence, but also because each member of an upper- and lower-case pair had been related to the same dictated sample. Thus, equivalence relations between experimental stimuli might have emerged anyway, even without the common locations of upper- and lower-case experimental stimuli within the teaching sequence.

It is noteworthy that the experimental variable, the conditional discriminations that were taught explicitly, may in principle be no different than the confounding variable, the sequence in which new letters were introduced. With respect to the experimental stimuli, dictated samples were related in common to two comparison stimuli; the same was true of stimulus locations within the teaching sequence. The dictated sample, the position within the teaching sequence, or both could have become equivalent to the pair of experimental stimuli to which they were related.

Therefore, even though Subject B.N.'s testimony pointed to a confounding variable, that variable would have produced equivalence relations no less genuine than those the experimental variable produced. With respect to upper- and lower-case control stimuli, the equivalence relations were not spurious, but the source of those relations was. They were produced by the "wrong" variable.

Although several writers (Dugdale & Lowe, 1990; Lowe, 1986; Stoddard &

McIlvane, 1986) have pointed out the need for caution in interpreting a subject's verbal reports in experiments on equivalence relations (an extreme example was Subject B.N.'s explanation, "I used my brain"), the accuracy of the subject's report in this instance is not really an issue. That report did identify a variable which, if not controlled, could influence experimental results. Subsequent experiments in our laboratory have therefore avoided consistency in the sequence of introduction of new stimuli when teaching subjects conditional discriminations as prerequisites for equivalence relations.

Other variables than those suggested by Subject B.N.'s remarks might also have produced the emergent relations between control stimuli. Let us suppose, for example, that upper- and lower-case experimental stimuli did become related because they had both become equivalent to the same dictated sample. Then, when faced for the first time (in a test) with a control stimulus as the sample and the five letters of the "other team" as comparisons, the subject might first have eliminated the three experimental comparison stimuli by exclusion (Dixon, 1977; McIlvane & Stoddard, 1981, 1985). The subject might then have arbitrarily selected one of the control stimuli just to get on with the next trial. Support for this possibility comes from Table 2; all subjects distributed their choices of comparisons differentially, depending on whether the samples were experimental or control stimuli. These differential distributions of choices suggest the formation of two stimulus classes, experimental and control.

For example, suppose the sample was the control stimulus, upper-case omega. Because some of the comparisons, the lower-case experimental stimuli xi, gamma, and lambda, already "went to" specified members of the other team, the child might have eliminated them as possible choices and arbitrarily selected one of the uncommitted control stimuli, say lower-case sigma. On a subsequent trial, the other upper-case control stimulus, phi, would have been presented as the sample. This time, the child might have excluded not only the experimental stimuli as possible choices, but lower-case sigma also, because that letter had already been selected in relation to upper-case omega.

Having learned to remain consistent, the child might have continued to select sigma or delta whenever omega or phi, respectively, appeared as the sample. These arbitrary relations between control stimuli would then have needed only symmetry to make them look like indicators of equivalence. If the subject's comparison selections were actually controlled by a larger pattern—a compound consisting of the two related stimuli—even symmetry need not be assumed.

Other factors, too, can contribute to the emergence of untaught conditional discriminations that seem to constitute positive evidence of equivalence, but do so only through experimental artifact (Harrison & Green,1990; Sidman, 1987). New kinds of procedural artifacts are constantly being discovered. Particularly deceptive are instances in which equivalence arises from sources other than those the experimenter or theorist supposes to have been causal. When known or suspected, such factors can be controlled. Where they have not been controlled, a detailed account of procedures can permit others to evaluate their possible influence upon experimental results, thereby increasing the validity and utility of descriptive and theoretical accounts of stimulus equivalence that are held to be supported by or derived from those results.

REFERENCES

Barnes, T. (1990). *Equivalence without symmetry? A stimulus artifact.* Unpublished M.A. thesis, *Northeastern University*, Boston.

Bush, K. M., Sidman, M., & de Rose, T. (1989). Contextual control of emergent equivalence relations. *Journal of the Experimental Analysis of Behavior, 51*, 29-45.

Carrigan, P. F. (1986). *Conditional discrimination and transitive relations: A theoretical and experimental analysis.* Unpublished Ph.D. dissertation, *Northeastern University*, Boston.

Carter, D. E., & Werner, T. J. (1978). Complex learning and information processing by pigeons: A critical analysis. *Journal of the Experimental Analysis of Behavior, 29*, 565-601.

Constantine, B. (1981). *An experimental analysis of stimulus control in simple conditional discriminations.* Unpublished Ph.D. dissertation, *Northeastern University*, Boston.

Dixon, L. S. (1977). The nature of control by spoken words over visual stimulus selection. *Journal of the Experimental Analysis of Behavior, 27*, 433-442.

Dugdale, N., & Lowe, C. F. (1990). Naming and stimulus equivalence. In D. E. Blackman & H. Lejeune (Eds.), *Behaviour analysis in theory and practice: Contributions and controversies* (pp. 115-138). Hillsdale, NJ: Erlbaum.

Harrison, R.J., & Green, G. (1990). Development of conditional and equivalence relations without differential consequences. *Journal of the Experimental Analysis of Behavior, 54*, 225-237.

Iversen, I. H., Sidman, M., & Carrigan, P. (1986). Stimulus definition in conditional discriminations. *Journal of the Experimental Analysis of Behavior, 45*, 297-304.

Johnson, C., & Sidman, M. (1990, May). *Stimulus classes established by sample-and-S- conditioninal-discrimination training.* Poster presented at the 16th annual meeting of the Association for Behavior Analysis, Nashville, TN.

Lazar, R. M., Davis-Lang, D., & Sanchez L. (1984). The formation of visual stimulus equivalences in children. *Journal of the Experimental Analysis of Behavior, 41*, 251-266.

Lowe, C. F. (1986, May). *The role of verbal behavior in the emergence of equivalence relations.* Paper presented at the 12th annual meeting of the Association for Behavior Analysis, Milwaukee, WI.

McIlvane, W. J., & Stoddard, L. T. (1981). Acquisition of matching-to-sample performances in severe retardation: Learning by exclusion. *Journal of Mental Deficiency Research, 25*, 33-48.

McIlvane, W. J., & Stoddard, L. T. (1985). Complex stimulus relations and exclusion in mental retardation. *Analysis and Intervention in Developmental Disabilities, 5*, 307-321.

Saunders, R. R., Saunders, K. J., Kirby, K. C., & Spradlin, J. E. (1988). The merger and development of equivalence classes by unreinforced conditional selection of comparison stimuli. *Journal of the Experimental Analysis of Behavior, 29*, 145-162.

Saunders, R. R., Wachter, J., & Spradlin, J. E. (1988). Establishing auditory stimulus control over an eight-member equivalence class via conditional discrimination procedures. *Journal of the Experimental Analysis of Behavior, 49*, 95-115.

Sidman, M. (1971). Reading and auditory-visual equivalences. *Journal of Speech and Hearing Research, 14*, 5-13.

Sidman, M. (1980). A note on the measurement of conditional discrimination. *Journal of the Experimental Analysis of Behavior, 33*, 285-289.

Sidman, M. (1986). Functional analysis of emergent verbal classes. In T. Thompson & M. D. Zeiler (Eds.), *Analysis and integration of behavioral units* (pp. 213-245). Hillsdale, NJ: Erlbaum.

Sidman, M. (1987). Two choices are not enough. *Behavior Analysis, 22*, 11-18.

Sidman, M. (in press). Equivalence relations: Some basic considerations. In L.J. Hayes & S.C. Hayes (Eds.), *Dialogues on verbal behavior: Proceedings of the third International Institute on Verbal Relations.* Reno, NV: Context Press.

Sidman, M., Cresson, O., Jr., & Willson-Morris, M. (1974). Acquisition of matching to sample via mediated transfer. *Journal of the Experimental Analysis of Behavior, 22*, 261-273.

Sidman, M., Kirk, B., & Willson-Morris, M. (1985). Six-member stimulus classes generated by conditional-discrimination procedures. *Journal of the Experimental Analysis of Behavior, 43*, 21-42.

Sidman, M., & Tailby, W. (1982). Conditional discrimination vs. matching to sample: An expansion of the testing paradigm. *Journal of the Experimental Analysis of Behavior, 37*, 5-22.

Sidman, M., Willson-Morris, M., & Kirk, B. (1986). Matching-to-sample procedures and the development of equivalence relations: The role of naming. *Analysis and Intervention in Developmental Disabilities, 6*, 1-19.

Sigurdardottir, Z. G., Green, G., & Saunders, R. R. (1990). Equivalence classes generated by sequence training. *Journal of the Experimental Analysis of Behavior, 53*, 47-63.

Spradlin, J. E., Cotter, V. W., & Baxley, N. (1973). Establishing a conditional discrimination without direct training: A study of transfer with retarded adolescents. *American Journal of Mental Deficiency, 77*, 556-566.

Stoddard, L. T., & McIlvane, W. J. (1986). Stimulus control research and developmentally disabled individuals. *Analysis and Intervention in Developmental Disabilities, 6*, 155-178.

Wulfert, E., & Hayes, S. C. (1988). Transfer of a conditional ordering response through conditional equivalence classes. *Journal of the Experimental Analysis of Behavior, 50*, 125-144.